Data-Driven Safety Quantification using Infinite-Dimensional Robust Convex Optimization

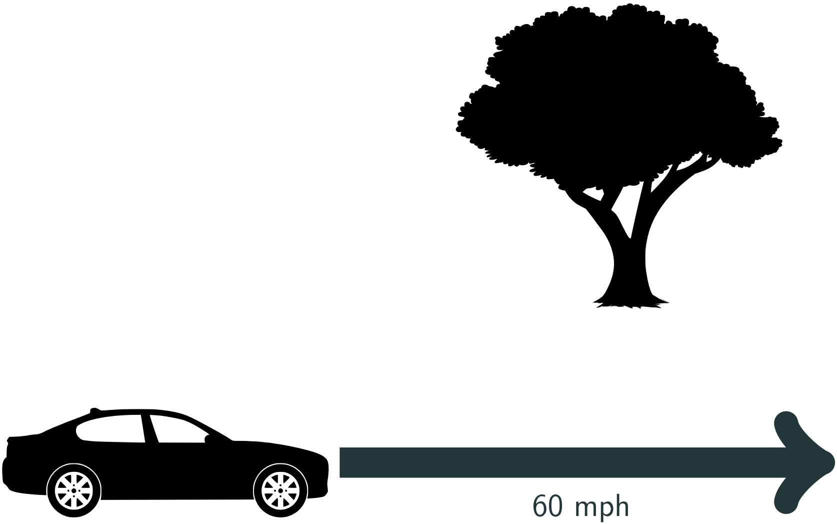
Jared Miller, Mario Sznaier

MSL Lab Meeting

May 19, 2023

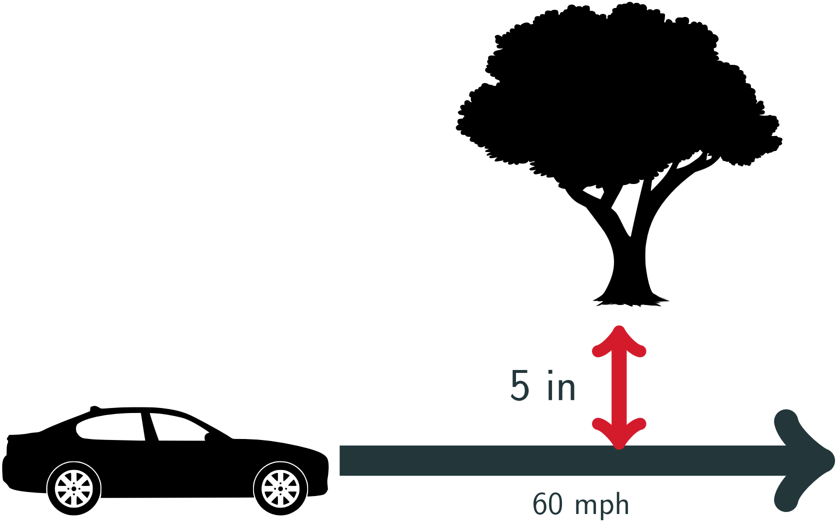
Northeastern University

Safety Example



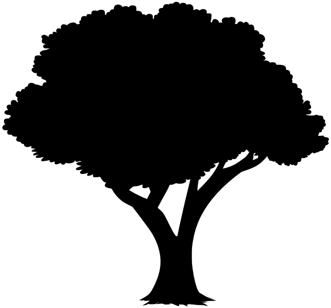
Safety Example (Barrier/Density Function)

Safety Example (Distance Estimate)



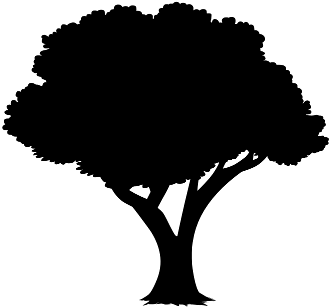
Safety Example (Distance Estimate)

Safety Quantification



Safety Example (Crash Control Effort)

Safety Quantification



Motivation: Epidemic

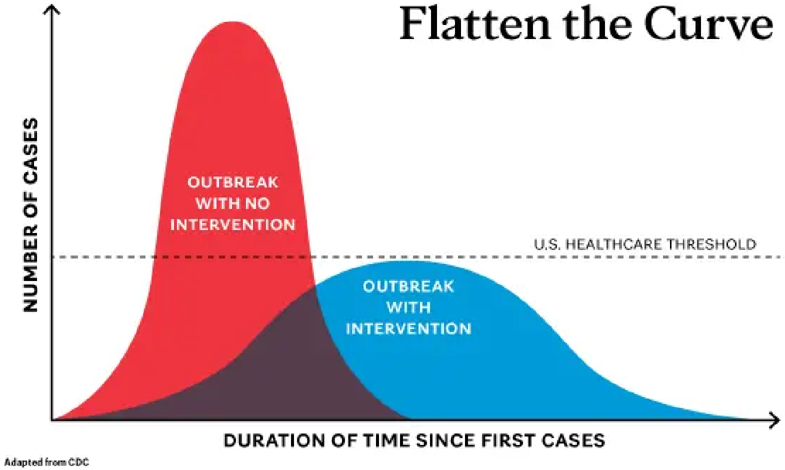
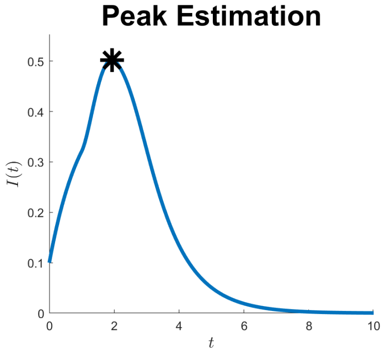
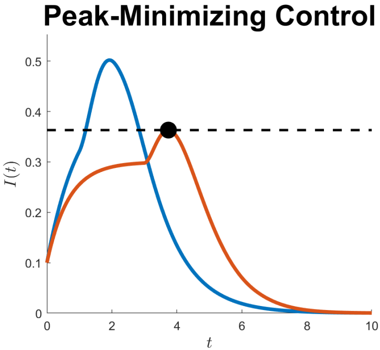


Image credit to Mayo Clinic News Network

Problems Covered





Flow of Presentation

Review peak estimation problem and SOS methods

Observe common pattern (peak, distance, crash):

- Input-affine dynamics
- Semidefinite-representable uncertainty

Use structure to simplify Lie derivative constraint

Apply method data-driven systems analysis

Peak and Sum-of-Squares Background

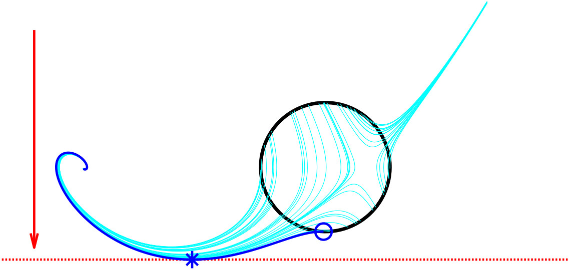
Peak Estimation Problem

Find maximum value of $p(x)$ along trajectories

$$P^* = \sup_{t,\ x_0 \in X_0} p(x(t))$$

$$\dot{x}(t) = f(t, x(t)) \qquad\qquad t \in [0, T]$$

$$x(0) = x_0 \in X_0$$



$$\dot{x} = [x_2, -x_1 - x_2 + x_1^3/3]$$

Peak Function Program

Infinite dimensional linear program (Cho, Stockbridge, 2002)

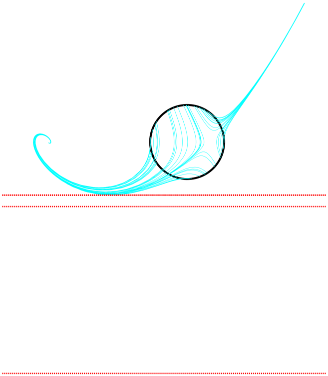
Uses auxiliary function $v(t, x)$

$$
\begin{aligned}
d^* = \inf_{\gamma \in \mathbb{R}} \quad & \gamma && && \text{(1a)} \\
& \gamma \geq v(0, x) && \forall x \in X_0 && \text{(1b)} \\
& \mathcal{L}_f v(t, x) \leq 0 && \forall (t, x) \in [0, T] \times X && \text{(1c)} \\
& v(t, x) \geq p(x) && \forall (t, x) \in [0, T] \times X && \text{(1d)} \\
& v \in C^1([0, T] \times X) && && \text{(1e)}
\end{aligned}
$$

Lie Derivative $\mathcal{L}_f v(t, x) = \partial_t v + f(t, x) \cdot \nabla_x v$

$P^* = d^*$ holds if $[0, T] \times X$ is compact, f Lipschitz

Peak Estimation Example Bounds



Converging bounds to min. $x_2 = -0.5734$ (moment-SOS)

Box region $X = [-2.5, 2.5]$, time $t \in [0, 5]$

Max. PSD size: $\binom{(n+1)+(d+\lfloor \deg f/2 \rfloor)}{n+1}$ (Fantuzzi, Goluskin, 2020)

Sum-of-Squares Method

Every $c \in \mathbb{R}$ satisfies $c^2 \geq 0$

Sufficient: $q(x) \in \mathbb{R}[x]$ nonnegative if $q(x) = \sum_i q_i^2(x)$

Exists $v(x) \in \mathbb{R}[x]^s$, *Gram* matrix $Z \in \mathbb{S}_+^s$ with $q = v^T Z v$

Sum-of-Squares (SOS) cone $\Sigma[x]$

$$
\begin{aligned}
&x^2y^4 - 6x^2y^2 + 10x^2 + 2xy^2 + 4xy - 6x + 4y^2 + 1 \\
=&(x + 2y)^2 + (3x - 1 - xy^2)^2
\end{aligned}
$$

Motzkin Counterexample (nonnegative but not SOS)

$$x^2y^4 + x^4y^2 - x^2y^2 + 1$$

Sum-of-Squares Method (cont.)

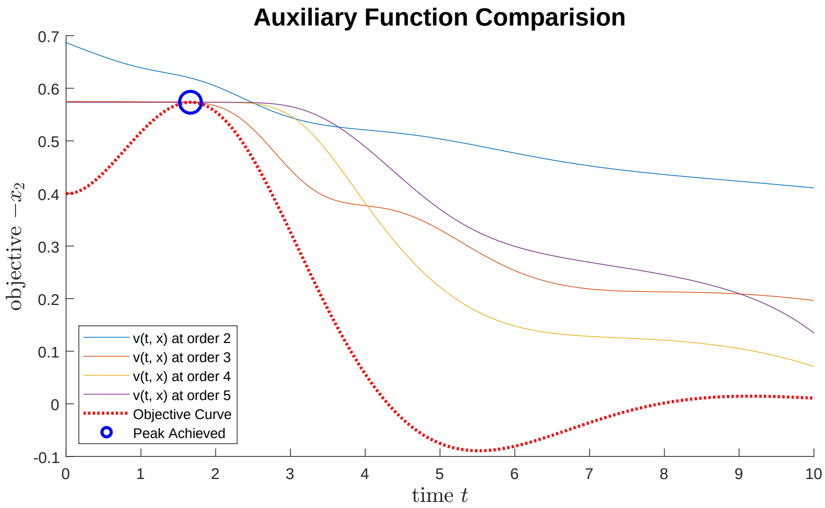
Putinar Positivestellensatz (Psatz) nonnegativity certificate over set $\mathbb{K} = \{x \mid g_i(x) \geq 0, h_j(x) = 0\}$:

$$q(x) = \sigma_0(x) + \textstyle\sum_i \sigma_i(x) g_i(x) + \sum_j \phi_j(x) h_j(x)$$
$$\exists \sigma_0(x) \in \Sigma[x], \quad \sigma_i(x) \in \Sigma[x], \quad \phi_j \in \mathbb{R}[x]$$

Psatz at degree $2d$ is an SDP, monomial basis: $s = \binom{n+d}{d}$

Archimedean: $\exists R \geq 0$ where $R - \|x\|_2^2$ has Psatz over $\mathbb{K}$

Auxiliary Evaluation along Optimal Trajectory



Optimal $v(t,x)$ should be constant until peak is achieved

Peak Estimation with Uncertainty

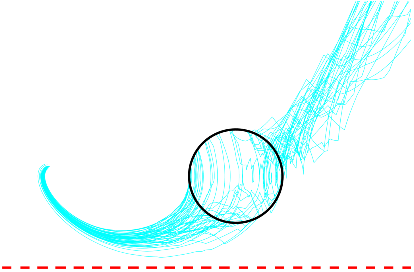
Dynamics $\dot{x} = f(t, x(t), w(t))$

Uncertain process $w(t) \in W, \ \forall t \in [0, T]$

Time-dependent $w(\cdot)$ with no continuity assumptions

$$\mathcal{L}_{f(t,x,w)} v(t,x) \leq 0 \qquad \forall (t, x, w) \in \forall [0, T] \times X \times W$$

System with Uncertainty Example

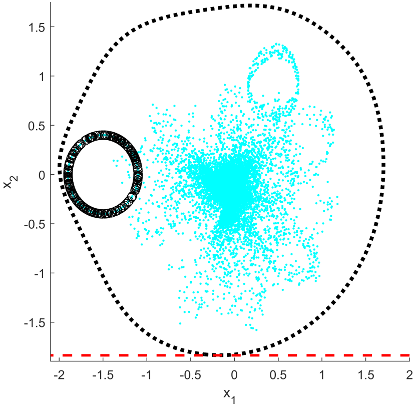


$$\dot{x}(t) = [x_2(t), -x_1 w(t) - x_2(t) + x_1(t)^3/3]$$
$$w(t) \in [0.5, 1.5], \text{ no continuity}$$

(Miller, Henrion, Sznaier, Korda, 2021)

Other types of Uncertainty Structures

- Switching Uncertainty
- Polytoptic Restriction
- Slew-Rate Bounded
- Discrete-Time
- Stochastic



Discrete dynamics with switching and time-dependent uncertainty

Robust Counterparts

Assumptions

Set $[0, T] \times X$ is compact

Uncertainty W is compact and convex

Dynamics $f(t, x, w)$ are Lipschitz in $[0, T] \times X$

Input-affine $f(t, x, w) = f_0(t, x) + \sum_{\ell=1}^{L} w_\ell f_\ell(t, x)$

Robust Counterpart Example: Box

(Ben-Tal, Nemirovskii, "Robust Optimization" 2009)

Original β-feasible problem with unknown $\|w\|_\infty \leq 1$

$$\forall w: \quad a_0^T\beta + \sum_{\ell=1}^L w_\ell a_\ell^T \beta \leq b_0 + \sum_{\ell=1}^L w_\ell b_\ell \tag{2}$$

Equivalent program with w eliminated:

$$\max_{\|w\|_\infty \leq 1} \left(\sum_{\ell=1}^L w_\ell [a_\ell^T \beta - b] \right) \leq b_0 - a_0^T \beta \tag{3a}$$

$$\sum_{\ell=1}^L |a_\ell^T \beta - b| \leq b_0 - a_0^T \beta \tag{3b}$$

Semidefinite-Representable (SDR) Set

Build up uncertainty set W using

K_s	Cone
λ_s	Lifting variable
(A_s, G_s, e_s)	Constraint description

Form the intersection

$$W = \cap_s \{\exists \lambda_s \in \mathbb{R}^{q_s} : A_s w + G_s \lambda_s + e_s \in K_s\}$$

SDR: All $K_s \subseteq$ PSD cone (projections of spectahedra)

Robust Counterparts (General)

Original problem with SDR uncertainty

$$\forall w \in W : \qquad a_0^T \beta + \sum_{\ell=1}^{L} w_\ell a_\ell^T \beta \leq b_0 + \sum_{\ell=1}^{L} w_\ell b_\ell$$

Robust counterpart (sufficient condition)

$$\begin{aligned}
&\textstyle\sum_{s=1}^{N_s} e_s^T \zeta_s + a_0^T \beta \leq b_0 && \\
&G_s^T \zeta_s = 0 && \forall s = 1..N_s \\
&\textstyle\sum_{s=1}^{N_s} (A_s^T \zeta_s)_\ell + a_\ell^T \beta = b_\ell && \forall \ell = 1..L \\
&\zeta_s \in K_s^* && \forall s = 1..N_s.
\end{aligned}$$

Nonconservative if (K convex, pointed, non-polyhedral Slater)

Robust Counterparts for Lie Constraint

Original **strict** constraint

$$\mathcal{L}_f v(t,x,w) < 0 \qquad \forall (t,x,w) \in [0,T] \times X \times W$$

Specific with $a_0, a_\ell = 0,\ b_0 = \mathcal{L}_{f_0} v,\ b_\ell = f_\ell \cdot \nabla_x v$

Robust counterpart with multipliers ζ

$$\begin{aligned}
&\mathcal{L}_{f_0} v(t,x) + \textstyle\sum_{s=1}^{N_s} e_s^T \zeta_s(t,x) < 0 && \forall [0,T] \times X \\
&G_s^T \zeta_s(t,x) = 0 && \forall s = 1..N_s \\
&\textstyle\sum_{s=1}^{N_s} (A_s^T \zeta_s(t,x))_\ell + f_\ell(t,x) \cdot \nabla_x v(t,x) = 0 && \forall \ell = 1..L \\
&\zeta_s(t,x) \in K_s^* && \forall s = 1..N_s
\end{aligned}$$

Conditions for Nonconservatism

Strict robust Lie constraint nonconservative if:

1. K_s convex, pointed, non-polyhedral Slater
2. $[0, T] \times X$ is compact
3. (A_s, G_s) constant in (t, x)
4. (f_0, f_ℓ, e_s) continuous in (t, x)

Peak Decomposed Program

Example: Polytopic uncertainty $W = \{w \mid Aw \leq b\}$

Only the Lie Derivative constraint changes

$$
\begin{aligned}
d^* = \ & \min_{\gamma \in \mathbb{R}} \gamma \\
& \gamma \geq v(0, x) & & \forall x \in X_0 \\
& \mathcal{L}_{f_0} v(t, x) + b^T \zeta(t, x) \leq 0 & & \forall (t, x) \in [0, T] \times X \\
& (A^T)_\ell \zeta(t, x) = (f_\ell \cdot \nabla_x) v(t, x) & & \forall \ell = 1..L \\
& v(t, x) \geq p(x) & & \forall (t, x) \in [0, T] \times X \\
& v(t, x) \in C^1([0, T] \times X) & & \\
& \zeta_k(t, x) \in C_+([0, T] \times X) & & \forall k = 1..m
\end{aligned}
$$

Applicable to any SDR W

Complexity and Input-Affine Structure

Assume input-affine $f(t, x, w) = f_0(t, x) + \sum_{\ell=1}^{L} w_\ell f_\ell(t, x)$

Size of largest PSD matrix in degree-d SDP:

$$\text{Original} \qquad \binom{1+n+L+d+\lceil \deg(f)/2 \rceil - 1}{1+n+L} = \binom{18}{13} = 8568$$

$$\text{Decomposed} \qquad \binom{1+n+d+\lceil \deg(f)/2 \rceil - 1}{1+N_x} = \binom{8}{3} = 56$$

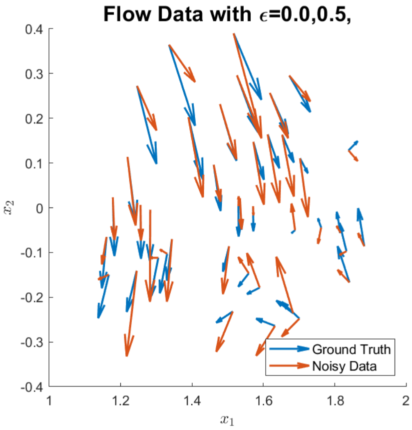
Values with $d = 4$, $L = 10$, $n = 2$

W polytope with 33 faces, 7534 vertices

Data-Driven Analysis

Sampling: Flow System

Data $\mathcal{D} = \{(t_j, x_j, \dot{x}_j)\}_j$ under mixed L_∞-bounded noise



$$\dot{x} = [x_2, -x_1 - x_2 + x_1^3/3]$$

Dynamics Model

Given data $\mathcal{D}$, budget ϵ, system model $\{f_0, f_\ell\}$

Parameterize ground truth F by functions in dictionary

$$\dot{x}(t) = f(t, x, w) = f_0(t, x) + \sum_{\ell=1}^{L} w_\ell f_\ell(t, x)$$

Ground truth satisfies corruption $J(w^*) \leq \epsilon$

L_∞ example: $J(w) = \max_j \|f(t_j, x_j, w) - \dot{x}_j\|_\infty$

Peak Estimation Examples

Elliptope Constraint on Input

$$W = \left\{ w \in \mathbb{R}^3 : \begin{bmatrix} 1 & w_1 & w_2 \\ w_1 & 1 & w_3 \\ w_2 & w_3 & 1 \end{bmatrix} \succeq 0 \right\}.$$

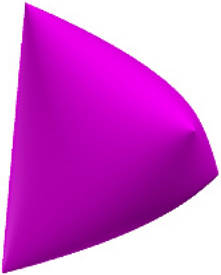
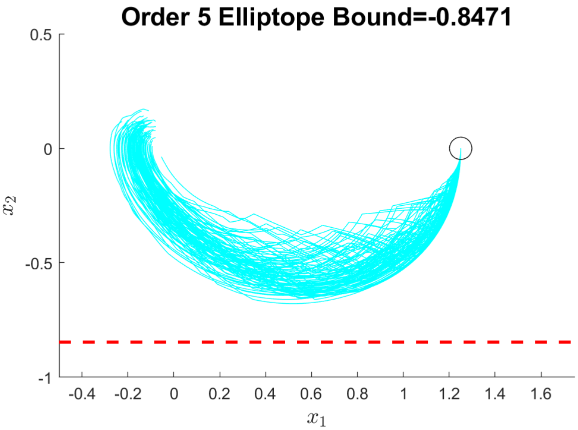


Image credit: Dattoro

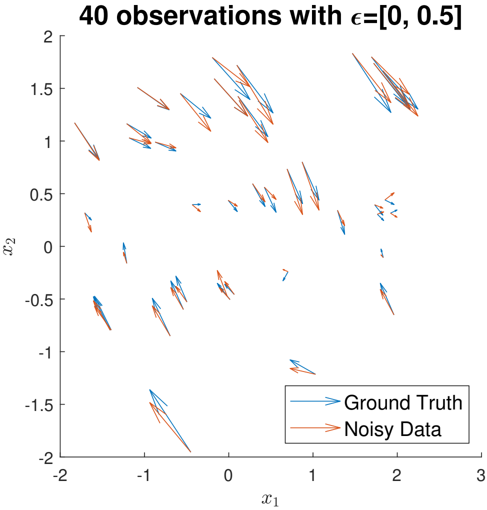
Peak Estimation Example (Flow-Elliptope)



$$f(t,x,w) = \begin{bmatrix} x_2 \\ -x_1 - x_2 + x_1^3/3 + w_1 x_1 + w_2 x_1 x_2 + w_3 x_3 \end{bmatrix}$$

Uses polynomial matrix inequalies of size 3×3

Peak Estimation Example (Flow)



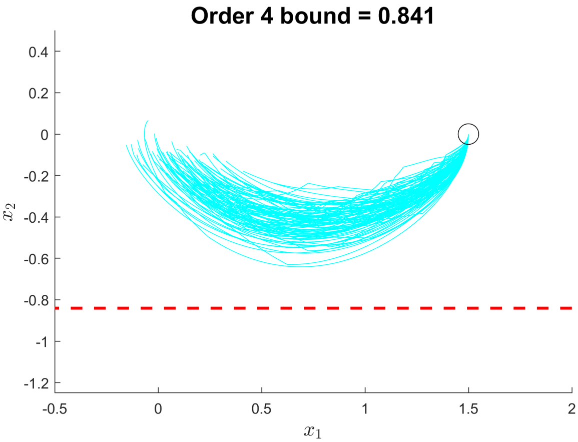
$$\dot{x} = [x_2,\ -x_1 - x_2 + x_1^3/3],\quad T = 5$$

Peak Estimation Example (Flow)

$$\dot{x} = [x_2, \ -wx_1 - x_2 + x_1^3/3]$$

$$L = 1, \ m = 80 \ (2 \text{ nonredundant})$$

Peak Estimation Example (Flow)



$$\dot{x} = [x_2,\ \text{cubic}(x_1, x_2)]$$

$$L = 10,\ \ m = 80\ (33\ \text{nonredundant})$$

Distance Estimation

Unsafe set X_u, point-set distance $c(x; X_u) = \inf_{y \in X_u} c(x, y)$

$$P^* = \sup_{t,\, x_0 \in X_0,\, w} -c(x(t \mid x_0); X_u)$$

$$\dot{x}(t) = f(t, x(t), w(t)) \qquad \forall t \in [0, T], \quad x(0) = x_0$$

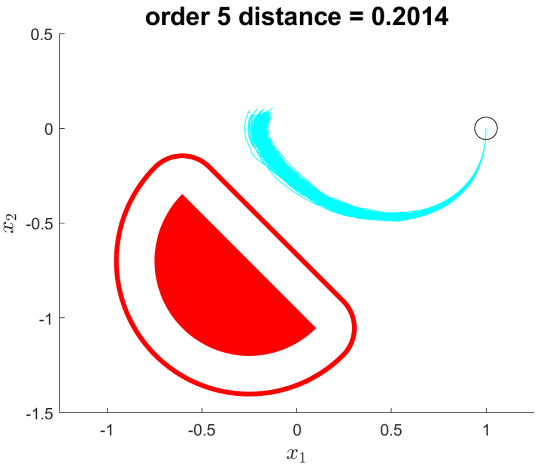
Polynomial-expressible constraints with $\phi(x)$

$$v(t, x) \geq \phi(x) \qquad \forall (t, x) \in [0, T] \times X$$

$$\phi(x) \geq -c(x, y) \qquad \forall (x, y) \in X \times X_u$$

No change to Lie derivative $\mathcal{L}_f v(t, x, w) \leq 0$

Distance Estimation Example (Flow)



$$\dot{x} = [x_2,\ \text{cubic}(x_1, x_2)]$$

$$L = 10,\ \ m = 80\ (33\ \text{nonredundant})$$

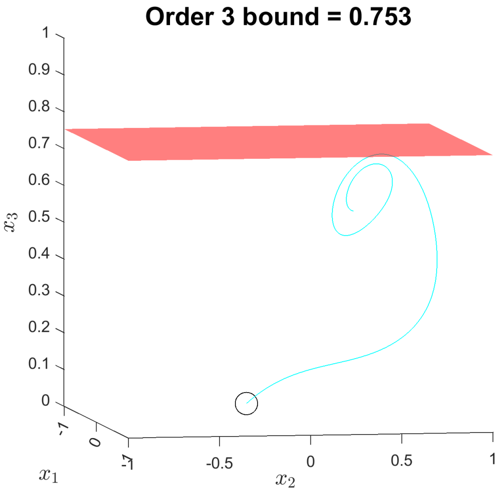
Peak Estimation Example (Twist)

Dynamics model:

$$\dot{x}_i = A_{ij}x_j - B_{ij}(4x_j^3 - 3x_j)$$

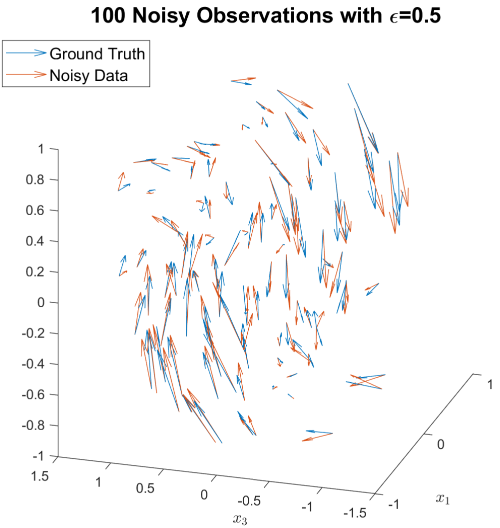
$$A = \begin{bmatrix} -1 & 1 & 1 \\ -1 & 0 & -1 \\ 0 & 1 & -2 \end{bmatrix}$$

$$B = \begin{bmatrix} -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$



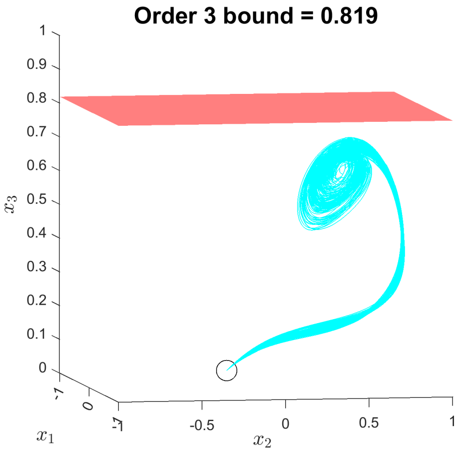
$$X_0 = [-1, 0, 0], \quad T = 8$$

Peak Estimation Example (Twist)



$$m = 2N_sN_x = 600 \text{ constraints}$$

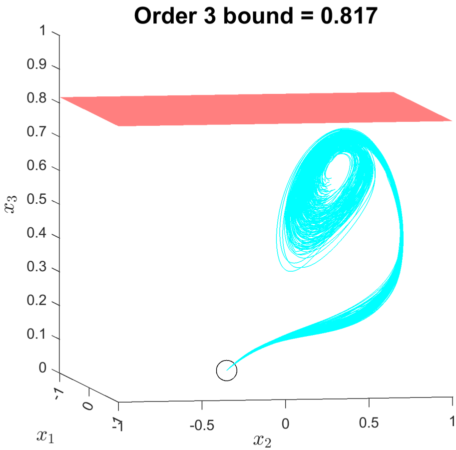
Peak Estimation Example (Twist)



Unknown A, Known B

$L = 9, \; m = 600$ (34 nonredundant)

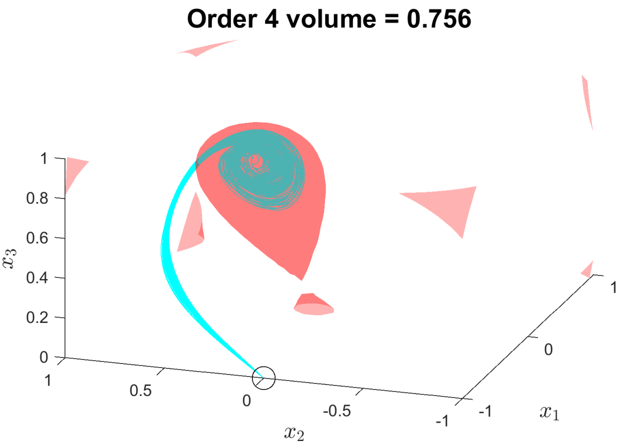
Peak Estimation Example (Twist)



Known A, Unknown B

$$L = 9, \ m = 600 \text{ (30 nonredundant)}$$

Reachable Set Estimation Example (Twist)



Unknown A, Known B

$L = 9,\ m = 600$ (34 nonredundant)

Crash-Safety

Crash-Safety

Corruption in L_∞-bounded setting

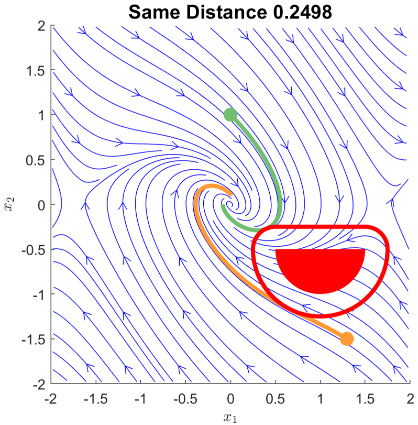
$$J(w) = \max_k \| f_0(t_k, x_k) + \textstyle\sum_{\ell=1}^{L} w_\ell f_\ell(t_k, x_k) - y_k \|_\infty$$
$$= \max(h - \Gamma w) \quad \text{for some polytope } \ (\Gamma, h)$$

How much data corruption is needed to crash?

$$Q^* = \inf_{t,\ x_0,\ w} \ \sup_{t' \in [0,t]} J(w(t'))$$
$$\dot{x}(t') = f(t', x(t'), w(t')) \qquad \forall t' \in [0, T]$$
$$x(t \mid x_0, w(\cdot)) \in X_u$$
$$w(\cdot) \in W,\ t \in [0, T],\ x_0 \in X_0$$

Example Crash-Bounds

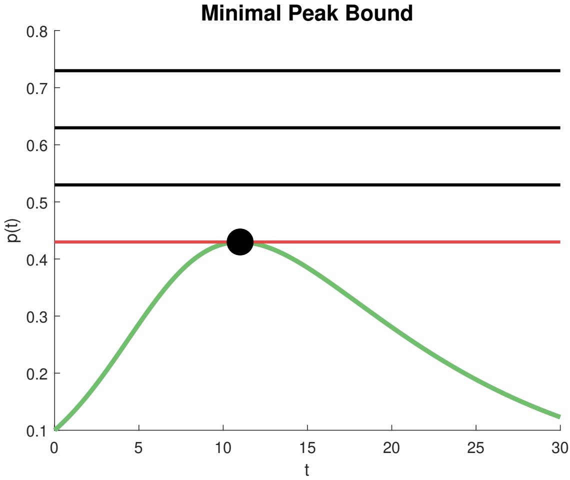
Two trajectories have same distance, different crash-bounds



Green-Top $Q^* = 0.316$, Yellow-Bottom $Q^* = 0.622$

Peak Minimizing Control

Find minimum bound on the maximum p value



Crash-safety is Peak Minimizing Control

Peak-Minimizing Control

Add state $\dot{z} = 0$ (Molina, Rapaport, Ramírez 2022)

$$
\begin{aligned}
Q_z^* = \inf_{t,\ x_0,\ z,\ w} \quad & z \\
& \dot{x}(t') = f(t', x(t'), w(t')) && \forall t' \in [0, T] \\
& \dot{z}(t') = 0 && \forall t' \in [0, T] \\
& J(w(t')) \leq z && \forall t' \in [0, T] \\
& x(t \mid x_0, w(\cdot)) \in X_u \\
& w(\cdot) \in W, \ \ t \in [0, T] \\
& x_0 \in X_0, z \in [0, J_{\max}]
\end{aligned}
$$

Drive down the z-upper-bound on $J(w)$

Crash-Bound Program

Consistency sets

$$Z = [0, J_{\max}] \qquad \Omega = \{(w, z) \in W \times Z : J(w) \leq z\}.$$

Optimal Control Problem with auxiliary $v(t, x, z) \in C^1$

$$
\begin{aligned}
d^* = \sup_{\gamma \in \mathbb{R},\ v} \quad & \gamma \\
& v(0, x, z) \geq \gamma && \forall (x, z) \in X_0 \times Z \\
& v(t, x, z) \leq z && \forall (t, x, z) \in [0, T] \times X_u \times Z \\
& \mathcal{L}_f v(t, x, z, w) \geq 0 && \forall (t, x, z, w) \in [0, T] \times X \times \Omega
\end{aligned}
$$

Crash Lie-decomposition

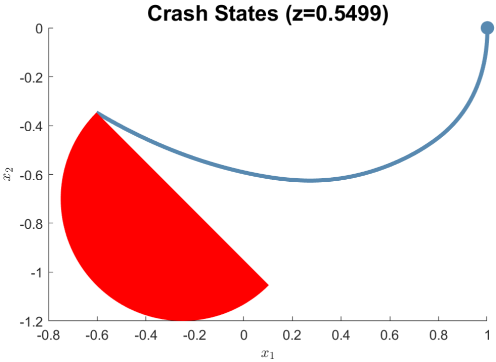
Exploit affine structure of $J(w) = \max_j (h - \Gamma w)_j$

Nonconservatively robustified Lie constraint

$$
\begin{aligned}
d^* = \sup_{\gamma \in \mathbb{R},\ v} \quad & \gamma \\
& v(0, x, z) \geq \gamma && \forall (x, z) \in X_0 \times Z \\
& v(t, x, z) \leq z && \forall (t, x, z) \in [0, T] \times X_u \times Z \\
& \mathcal{L}_{f_0} v - (z\mathbf{1} + h)^T \zeta \geq 0 && \forall (t, x, z) \in [0, T] \times X \times [0, J_{\max}] \\
& (\Gamma^T)_\ell \zeta + f_\ell \cdot \nabla_x v = 0 && \forall \ell = 1..L \\
& \zeta_j \in C_+([0, T] \times X \times Z) && \forall j = 1..2nT
\end{aligned}
$$

Data-Driven Flow Crash-Bound

CasADi trajectory matches SOS crash bound



Degree-4 crash bound also 0.5499

True $\epsilon = 0.5$, distance ≈ 0.2014

Flow Crash-Subvalue

Lower bound $q(x)$ for corruption needed to crash ($Q_{\max} < \infty$)

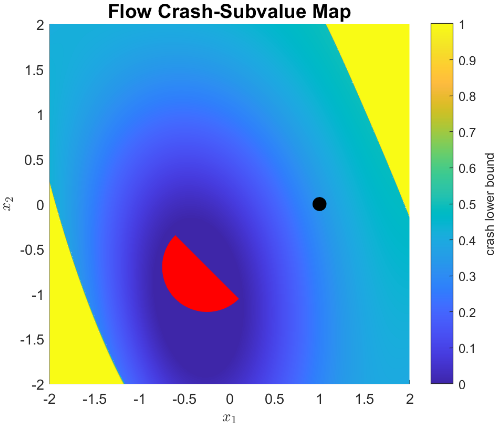
$$
\begin{aligned}
J^* = \sup \int_X q(x) dx & \\
v(0, x, z) \geq q(x) & \qquad \forall (x, z) \in X \times [0, Z_{\max}] \\
q(x) \leq Q_{\max} & \qquad \forall x \in X \\
z \geq v(t, x, z) & \qquad \forall (t, x, z) \in [0, T] \times X_u \times Z \\
\mathcal{L}_f v(t, x, z, w) \geq 0 & \qquad \forall (t, x, z, w) \in [0, T] \times X \times \Omega \\
v \in C^1([0, T] \times X \times Z) & \\
q \in C(X) &
\end{aligned}
$$

Based on Joint+Marginal optimization (Lasserre, 2010)

Flow Crash-Subvalue

Piecewise-polynomial subvalue

$q_{1:d}(x) = \max(-l_u(x), \max_{d' \in 1..d} q_{d'}(x))$



Bound of $0.3399 \leq 0.5499$, but valid everywhere in X

Take-aways

Conclusion

Tractable safety quantification problems

More SOS constraints in fewer variables

Data-driven estimates given semidefinite-bounded noise

Other applications

- Barrier Functions
- Maximum controlled invariant sets
- Hybrid systems
- Set-set (shape) distance estimation
- Distance-maximizing control
- Reachable set estimation

Safety is Important

Quantify using Peak Estimation

Extra Material

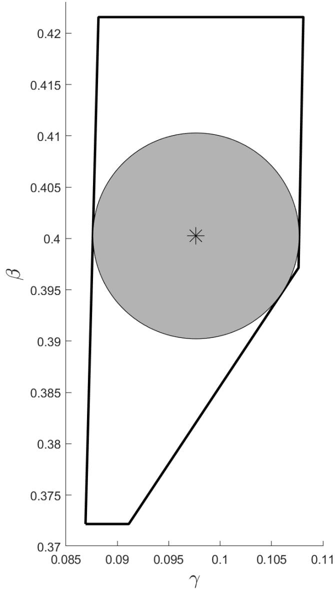
Preprocessing: Centering

Chebyshev center c: center of sphere with largest radius in W

Find through linear programming

$$\begin{aligned} &\max r \\ &A_k c + r\|A_k\|_2 \le b_k \qquad \forall k \\ &r \ge 0, c \in \mathbb{R}^L \end{aligned}$$

Shifted dynamics $f_0 \leftarrow f_0 + \sum_{\ell=1}^{L} c_\ell f_\ell$

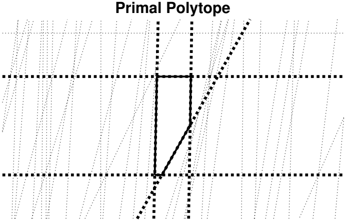


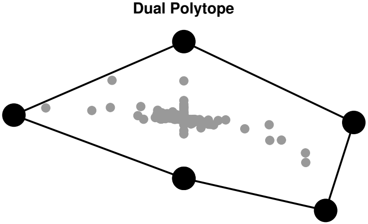
Preprocessing: Redundancy

Majority of $m = 2N_x N_s$ constraints are often redundant

Convex hull of dual polytope:
Time: $\Omega(m \log m + m^{\lfloor L/2 \rfloor})$

Linear program per constraint:
Time: $m \times \tilde{O}(mL + L^3)$
(Jan van den Brand *et. al.* 2020)





Polynomial Matrix Inequalities

SOS method (scalar): $q(x) \geq 0$

Extend to matrices $Q(x) \in \mathbb{S}^s_{++}$

SOS matrix: $Q(x) = R(x)^T R(x) \in \Sigma^s[x]$ for matrix $R(x)$

Gram matrix (PSD) constraint of size $s\binom{n+d}{d}$

Scherer Psatz: nonnegativity over constraint sets